

Goodness-of-Fit Test: Khmaladze Transformation vs Empirical Likelihood

Jiwoong Kim

Michigan State University

Abstract

This paper compares two asymptotic distribution free methods for goodness-of-fit test of one sample of location-scale family: Khmaladze transformation and empirical likelihood methods. The comparison is made from the perspective of empirical level and power obtained from simulations. When testing for normal and logistic null distributions, we try various alternative distributions and find that Khmaladze transformation method has better power in most cases. R-package which was used for the simulation is available online. See section 5 for the detail.

Keywords: Asymptotic distribution free; Khmaladze transformation; Empirical likelihood; Goodness-of-fit test

1 Introduction

A classical goodness-of-fit problem, i.e., the problem of testing whether a random sample comes from a specific distribution or from a given parametric family of distributions has been of interest to many fields for a long time. For example, the normality of sample has been commonly assumed in the vast literature of social and physical sciences. Since the final result will, then, heavily depend on normality assumption, goodness-of-fit test for normality has been a critical issue.

For goodness-of-fit test for general distributions, the various parametric and nonparametric tests have been proposed in the literature. The best known exemplary parametric and nonparametric tests are χ^2 test and Kolmogorov-Smirnov (K-S) test. The most attractive advantage of the K-S test is that asymptotic distribution of its test statistic under the

null hypothesis does not depend on the null distribution, when fitting a known continuous distribution. However, the K-S test loses this property when fitting a parametric family of distributions.

Seeking tests which preserve the desirable feature of being distribution free, we finally come up with two celebrated methods: Khmaladze martingale transformation (KMT) and empirical likelihood (EL). In this paper, we employ KMT and EL methods to test for a parametric location-scale family of distributions. Main goal is to compare these two methods and to show which method is superior. To that end, we report empirical levels and powers of KMT and EL methods. This paper is organized as follows. Section 2 briefly reviews KMT and EL methods. In section 3, we report our findings obtained from simulation.

2 KMT & EL methods

KMT method has not gained much attention despite asymptotic distribution free (ADF) property. Koul provided an excellent review of KMT method in the chapter 9 of Fan and Koul (2006). The review of KMT method in this paper has a root in his work. Let X_1, \dots, X_n be independent and identically distributed (i.i.d.) random sample of location-scale family where F is their common distribution function (d.f.), and f is absolutely continuous density. Assume that \dot{f} exists almost everywhere such that $0 < \int (\dot{f}/f)^2 dF < \infty$. Let μ and σ denote unknown location and scale parameters, respectively. Consider a problem to test

$$(2.1) \quad H_0 : F(x) = F_0((x - \mu)/\sigma), \quad \text{vs} \quad H_a : F(x) \neq F_0((x - \mu)/\sigma)$$

where F_0 is a known d.f. Define

$$(2.2) \quad \begin{aligned} Z_i &:= (X_i - \mu)/\sigma, & \hat{Z}_i &:= (X_i - \hat{\mu}_n)/\hat{\sigma}_n, \\ F_n(x) &:= n^{-1} \sum_{i=1}^n I(Z_i \leq x), & \hat{F}_n(x) &:= n^{-1} \sum_{i=1}^n I(\hat{Z}_i \leq x). \end{aligned}$$

where $\hat{\mu}_n$ and $\hat{\sigma}_n$ are consistent estimators of μ and σ under the null hypothesis.

As mentioned in the introduction, it is well known that the null distribution of classical K-S test based on F_n does not depend on F_0 . However, this fact does not hold any more

when the necessity of estimating μ and σ arises, i.e., a test based on \widehat{F}_n is not distribution free. Durbin (1973) showed that null distribution of the test based on \widehat{F}_n depends on the estimators of them as well as F_0 . To obtain ADF test, we pay attention to a martingale transformation based on \widehat{F}_n which was proposed by Khmaladze (1979, 1980). Let

$$(2.3) \quad \begin{aligned} \phi_0(x) &:= -\dot{f}_0(x)/f_0(x), \quad l(x) := (1, \phi_0(x), x\phi_0(x) - 1)', \\ p_0(t) &:= f_0(F_0^{-1}(t)), \quad q_0(t) := F_0^{-1}(t)f_0(F_0^{-1}(t)), \\ \Gamma_t &= \begin{pmatrix} 1-t & p_0(t) & q_0(t) \\ p_0(t) & \int_t^1 \dot{p}_0^2(u)du & \int_t^1 \dot{p}_0(u)\dot{q}_0(u)du \\ q_0(t) & \int_t^1 \dot{p}_0(u)\dot{q}_0(u)du & \int_t^1 \dot{q}_0^2(u)du \end{pmatrix}. \end{aligned}$$

Define martingale transformed process

$$(2.4) \quad \widehat{U}_n(t) := n^{-1/2} \sum_{i=1}^n \left\{ I(\widehat{Z}_i \leq z) - l(\widehat{Z}_i)' \int_{-\infty}^{z \wedge \widehat{Z}_i} \Gamma_{F_0^{-1}(x)}^{-1} l(x) dF_0(x) \right\}, \quad t = F_0(z), \quad z \in \mathbb{R}.$$

Then, weak convergence of \widehat{U}_n to Brownian motion in uniform metric follows from Khmaladze (1981): see the section 4 for the details. Hence, any test based on $T := \sup_{0 \leq t \leq 1} |\widehat{U}_n(t)|$ is ADF. When we test H_0 in (2.1) via KMT method, we shall use T for the test statistic.

When Y_1, Y_2, \dots, Y_n are i.i.d. observations from distribution K , EL is defined as

$$(2.5) \quad L(K) = \prod_{i=1}^n K(\{Y_i\}) = \prod_{i=1}^n p_i$$

where $p_i = K(\{Y_i\}) = K(Y_i) - K(Y_i-)$. It is well-known that empirical d.f. $K_n(x) := n^{-1} \sum_{i=1}^n I(Y_i \leq x)$ maximizes (2.5). Let EL ratio denote $R(K) := L(K)/L(K_n)$. Owen (1988, 1990) used $R(K)$ and constructed a nonparametric confidence region and test for the mean of Y . Let K_0 be a d.f. with mean μ_0 . He proposed

$$(2.6) \quad \mathbb{R}(\mu_0) = \sup \left\{ \prod_{i=1}^n np_i \mid p_i \geq 0, \sum_{i=1}^n p_i = 1, \sum_{i=1}^n p_i Y_i = \mu_0 \right\},$$

and showed that for $Y \sim K_0$

$$(2.7) \quad -2 \log \mathbb{R}(\mu_0) \longrightarrow \chi_{(1)}^2,$$

which is an analog of Wilks's (1938) theorem for nonparametric likelihood. Owen (1990) referred to $\mathbb{R}(\mu_0)$ as “profile” empirical likelihood ratio (PELR) since nuisance parameters are “profiled out.”

The PELR has been of interest to statisticians and extended to various settings. Consider the case where unknown K has θ , a parameter of d -dimension. Qin and Lawless (1994) assumed that there exists information about K and θ : there are $h \geq d$ “unbiased estimating functions,” that is, $g_1(Y, \theta), \dots, g_h(Y, \theta)$ where $E_K[g_j(Y, \theta)] = 0$ for $j = 1, 2, \dots, h$. For example, let $E(Y) = \theta$ and $E(Y^2) = \kappa(\theta)$ where $\kappa(\cdot)$ is a known function. Then, g_1 and g_2 can be written as $g_1(Y, \theta) = Y - \theta$ and $g_2(Y, \theta) = Y^2 - \kappa(\theta)$, respectively. With the unbiased estimating functions, they considered an analog of (2.6), i.e., a maximization problem

$$\max \prod_{i=1}^n n p_i \quad \text{subject to } p_i \geq 0, \sum_{i=1}^n p_i = 1, \sum_{i=1}^n p_i g_j(Y_i, \theta) = 0, \quad j = 1, 2, \dots, h.$$

Hence, they linked PELR with finitely many constraints. By the method of Lagrange multipliers, they obtained the maximum and defined profile empirical log-likelihood ratio (PELLR)

$$(2.8) \quad l_E(\theta) = \sum_{i=1}^n \log \left[1 + \sum_{j=1}^h \lambda_j g_j(Y_i, \theta) \right],$$

where λ_j is the Lagrange multiplier corresponding to the constraint $\sum_{i=1}^n p_i g_j(Y_i, \theta) = 0$. They showed λ_j 's are determined in terms of θ and under $H_0 : \theta = \theta_0$

$$(2.9) \quad 2l_E(\theta_0) - 2l_E(\hat{\theta}) \longrightarrow \chi_{(d)}^2,$$

where $\hat{\theta}$ minimizes $l_E(\theta)$.

In contrast to Qin and Lawless (1994), Peng and Schick (2013) considered PELR approach combined with infinitely many constraints (or unbiased estimating functions), i.e., the number of constraints increases as the sample size increases. Again, let Y be r.v. which comes from unknown distribution K . Consider testing $H_0 : K = K_0$ where K_0 is a known fixed d.f. Define $\varphi_h(x) := \sqrt{2} \cos(h\pi x)$, for $h = 1, 2, \dots$. Consequently, we have for all h

$$\int_{[0,1]} \varphi_h(x) dx = 0, \quad \int_{[0,1]} \varphi_h^2(x) dx = 1.$$

Note that when $H_0 : K = K_0$ is true, $K_0(Y)$ will be a uniform r.v., and hence, $E[\varphi_h(K_0(Y))] = 0$ and $E[\varphi_h(K_0(Y))]^2 = 1$. With infinitely many unbiased estimating functions, $\varphi_h \circ K_0$, $h = 1, 2, \dots$, they proposed an analog of (2.6)

$$(2.10) \quad \mathbb{R}_n(K_0) = \sup \left\{ \prod_{i=1}^n np_i : p_i \geq 0, \forall i, \sum_{i=1}^n p_i = 1, \sum_{j=1}^n p_j \varphi_h(K_0(Y_j)) = 0, h = 1, \dots, m_n \right\}$$

and showed that under H_0

$$-2 \log \mathbb{R}_n(K_0) \longrightarrow \chi_{(m_n)}^2,$$

where m_n and n tend to infinity and m_n^3/n tends to 0. They extended the result to testing H_0 where underlying distribution K has a unknown d -dimensional parameter θ . With efficient estimator $\hat{\theta}$ -e.g., maximum likelihood estimator (MLE)-they derived $\mathbb{R}_n(K_{\hat{\theta}})$ -the test statistic in (2.10) with K_0 replaced by $K_{\hat{\theta}}$ -and showed that under the null hypothesis

$$-2 \log \mathbb{R}_n(K_{\hat{\theta}}) \rightarrow \chi_{(m_n-d)}^2.$$

Since their approach is free from the question of how many constraints should be used, we use $-2 \log R_n(K_{\hat{\theta}})$ for the test statistic when we implement EL method to test H_0 in (2.1).

At this time, it should be mentioned that EL approach has one critical drawback when it is employed for hypothesis testing. Note that all the EL methods introduced in this section solve the maximization problem subject to finitely or infinitely many constraints (or unbiased estimating functions). Assume that r.v. Y comes from unknown distribution K . Consider testing $H_0 : K = K_0$ vs $H_a : K = K_a$ via the EL method. Let θ_0 and θ_a denote parameters associated with K_0 and K_a , respectively. When we use the test statistic in (2.8) ((2.10)), constraints corresponding to H_0 and H_a are $E_{K_0}[g(Y, \theta_0)] = 0$ and $E_{K_a}[g(Y, \theta_a)] = 0$. When K_0 are similar to K_a , we will therefore obtain the similar test statistics $l_E(\theta_0)$ and $l_E(\theta_a)$ (or $\mathbb{R}_n(K_0)$ and $\mathbb{R}_n(K_a)$). As a result, it is very likely to make a type II error when true d.f. of Y is K_a , i.e., EL method will have very poor power. Examples of such K_0 and K_a are; logistic and normal; and logistic and Student's t (STT) where degrees of freedom (df) is greater than or equal to 5. The poor power of EL method for these two cases are illustrated in the next section. See, e.g., Table 2 and 3.

3 Simulation study

Let F_i for $i = N, L$ denote d.f. of standard normal and logistic r.v., respectively. Note that

$$F_N(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \exp(-y^2/2) dy, \quad F_L(x) = \frac{1}{1 + e^{-x}}.$$

In this simulation study, we report the findings obtained from goodness-of-fit test for two location-scale distributions: $H_0 : F(x) = F_i((x - \mu)/\sigma)$, $i = N, L$. Table 1 reports critical values for KMT and EL methods. The critical values of KMT test are available at home-

α	KMT	EL1				EL2			
		$n=50$	100	200	500	$n=50$	100	200	500
0.05	2.24	5.99	7.81	9.49	12.59	7.81	9.49	11.07	14.07
0.01	2.81	9.21	11.34	13.28	16.81	11.34	13.28	15.09	18.48

Table 1: Critical value for KMT and EL

pages.ecs.vuw.ac.nz/ray/Brownian which is made by Dr. R. Brownrigg. For those of EL method, we consider two different df's: $m_{n1} = \lfloor n^{1/3} \rfloor - 2$ and $m_{n2} = \lfloor n^{1/3} \rfloor - 1$ where $\lfloor x \rfloor$ is the largest integer not greater than x . Let EL1 and EL2 denote EL methods with df of m_{n1} and one of m_{n2} , respectively. When we generate $n = 50, 100, 200$, and 500 samples from each null distribution, we use 2 and 5 for location and scale parameters. We repeat random sample generation 1,000 times and obtain KMT and EL test statistics. In order to obtain two different EL test statistics, we use $\lfloor n^{1/3} \rfloor$ and $\lfloor n^{1/3} \rfloor + 1$ unbiased estimating functions for the constraints in (2.10). Empirical levels and powers are then calculated from dividing the number of rejection of null hypothesis by 1,000. We use MLE for $\hat{\mu}$ and $\hat{\sigma}$ in the subsequent sections.

3.1 Testing for normal distribution

In this section, we compare KMT and EL methods to test for normality. Let F denote standard normal d.f. and f be its density. $l(x)$ and $\Gamma_{F(x)}$ in (2.3) turn out

$$l(x) = (1, x, x^2 - 1)',$$

$$\Gamma_{F(x)} = \begin{pmatrix} 1 - F(x) & f(x) & xf(x) \\ f(x) & xf(x) + (1 - F(x)) & (1 + x^2)f(x) \\ xf(x) & (1 + x^2)f(x) & (x^3 + x)f(x) + 2(1 - F(x)) \end{pmatrix}.$$

Let

$$A_{11}(x) := 1 - F(x), \quad A_{12}(x) := (f(x), xf(x)), \quad A_{21}(x) := (f(x), xf(x))',$$

$$A_{22}(x) := \begin{pmatrix} xf(x) + (1 - F(x)) & (1 + x^2)f(x) \\ (1 + x^2)f(x) & (x + x^3)f(x) + 2(1 - F(x)) \end{pmatrix}.$$

Then the inverse of $\Gamma_{F(x)}$ also can be expressed in partitioned form, i.e.,

$$\Gamma_{F(x)}^{-1} = \begin{pmatrix} B_{11}(x) & B_{12}(x) \\ B_{21}(x) & B_{22}(x) \end{pmatrix}$$

where

$$(3.1) \quad \begin{aligned} B_{11}(x) &= (A_{11}(x) - A_{12}(x)A_{22}(x)^{-1}A_{21}(x))^{-1}, \\ B_{12}(x) &= -B_{11}(x)A_{12}(x)A_{22}(x)^{-1}, \quad B_{21}(x) = B_{12}(x)', \\ B_{22}(x) &= A_{22}(x)^{-1} + A_{22}(x)^{-1}A_{21}(x)B_{11}(x)A_{12}(x)A_{22}(x)^{-1}. \end{aligned}$$

Let $\tilde{F}(x) := 1 - F(x)$, $c_1(x) := \{-(x^2 + 1)f^2(x) + (x^3 + 3x)f(x)\tilde{F}(x) + 2\tilde{F}^2(x)\}^{-1}$ and $c_2(x) := 2\tilde{F}^3(x) + (x^3 + 3x)f(x)\tilde{F}^2(x) - (2x^2 + 3)f^2(x)\tilde{F}(x) + xf^3(x)$. Note that with $r(x) := (x, (x^2 - 1))'$, we have

$$\Gamma_{F(x)}^{-1}l(x)f(x) = \begin{bmatrix} f(x)B_{11}(x) + f(x)B_{12}(x)r(x) \\ f(x)B_{21}(x) + f(x)B_{22}(x)r(x) \end{bmatrix}.$$

Finally, with \hat{Z}_i in (2.2), we have

$$\begin{aligned}
(3.2) \quad & l(\widehat{Z}_i)' \Gamma_{F(x)}^{-1} l(x) f(x) \\
&= 2f(x)c_2(x) \left[\tilde{F}^2(x) + xf(x)\tilde{F}(x) - f^2(x) \right] \\
&\quad + \widehat{Z}_i f(x)c_1(x)c_2(x) \left[4x\tilde{F}^4(x) + (2x^4 + 8x^2 - 2)f(x)\tilde{F}^3(x) + (x^5 - 7x)f^2(x)\tilde{F}^2(x) \right. \\
&\quad \left. - (2x^4 + 3x^2 - 1)f^3(x)\tilde{F}(x) + (x^3 + x)f^4(x) \right] \\
&\quad + (\widehat{Z}_i^2 - 1)f(x)c_1(x)c_2(x) \left[2(x^2 - 1)\tilde{F}^4(x) + (x^5 + 2x^3 - 9x)f(x)\tilde{F}^3(x) \right. \\
&\quad \left. - (4x^4 + 9x^2 - 5)f^2(x)\tilde{F}^2(x) + (5x^3 + 9x)f^3(x)\tilde{F}(x) - 2(x^2 + 1)f^4(x) \right].
\end{aligned}$$

We then replace $l(\widehat{Z}_i)' \Gamma_{F(x)}^{-1} l(x) f(x)$ in (2.4) by (3.2).

Table 2 reports empirical levels and powers of KMT and EL methods. The first (second) columns of KMT and EL's represent those when α is 0.05 (0.01). The value corresponding to normal F represents the empirical level while those corresponding to others-logistic, STT with df of 5, mixture of two normal distributions (MTN), Cauchy, and Laplace-represent the powers. For the MTN, we use $0.9N(2, 5^2) + 0.1N(2, 15^2)$; for logistic, Cauchy, and Laplace distributions, we use 2 and 5 for location and scale parameters. (*) implies the corresponding method shows the closest empirical level to the α or highest power; e.g., KMT has the closest empirical level (0.01) and the highest power (0.152) when $n = 500$ with $\alpha = 0.01$ and $n = 50$ with $\alpha = 0.05$, respectively. It is hard to judge the superiority of two methods by the empirical level since neither KMT nor EL demonstrates better performance consistently.

However, it is not hard to tell which of the two methods is better in terms of the power. For all distributions except Laplace, KMT outperforms EL: KMT shows better power than EL in most n 's. When F is logistic, EL displays poor power (less than 0.4) as stated in the previous section. The power of KMT is more than or equal to almost twice that of EL when n is 200 or 500. For STT, MTN, and Cauchy, KMT still maintains superiority over EL even though that is weakened. As shown in the table, the performance gap between KMT and EL is widened when α is 0.01.

In contrast, Laplace F shows the opposite result: EL outperforms KMT for all n 's and α 's. But the superiority of EL over KMT is not as strong as that of KMT over EL in the

F	n	KMT	EL1	EL2	KMT	EL1	EL2
normal	50	0.026	0.088	0.058 (*)	0.012 (*)	0.020	0.022
	100	0.032	0.060 (*)	0.082	0.022	0.016 (*)	0.018
	200	0.032	0.044	0.050 (*)	0.016	0.006	0.008 (*)
	500	0.024	0.062	0.060 (*)	0.010 (*)	0.018	0.016
logistic	50	0.152 (*)	0.148	0.100	0.092 (*)	0.036	0.032
	100	0.200 (*)	0.148	0.168	0.120 (*)	0.056	0.032
	200	0.320 (*)	0.176	0.168	0.244 (*)	0.052	0.056
	500	0.632 (*)	0.372	0.384	0.532 (*)	0.160	0.160
STT	50	0.244	0.276 (*)	0.224	0.208 (*)	0.116	0.092
	100	0.400 (*)	0.304	0.264	0.292 (*)	0.132	0.136
	200	0.556 (*)	0.500	0.468	0.480 (*)	0.252	0.680
	500	0.884 (*)	0.868	0.848	0.804 (*)	0.680	0.660
MTN	50	0.382	0.410 (*)	0.346	0.352 (*)	0.224	0.164
	100	0.562 (*)	0.528	0.462	0.516 (*)	0.296	0.258
	200	0.812 (*)	0.756	0.706	0.716 (*)	0.566	0.532
	500	0.982 (*)	0.966	0.958	0.966 (*)	0.908	0.886
Cauchy	50	0.832	0.996 (*)	0.980	0.744	0.984 (*)	0.968
	100	0.988 (*)	0.984	0.968	0.948	0.984 (*)	0.968
	200	1.000 (*)	0.976	0.968	1.000 (*)	0.980	0.964
	500	1.000 (*)	0.990	0.964	1.000 (*)	0.970	0.960
Laplace	50	0.276	0.468 (*)	0.424	0.186	0.188 (*)	0.188 (*)
	100	0.480	0.788 (*)	0.676	0.380	0.516 (*)	0.444
	200	0.768	0.932 (*)	0.920	0.632	0.796 (*)	0.756
	500	0.980	1.000 (*)	1.000 (*)	0.964	1.000 (*)	1.000 (*)

Table 2: Empirical level and power obtained from testing $H_0 : F = F_N$

logistic case; when n reaches 500, the difference of powers between EL and KMT is less than 0.05 while counterpart of the logistic case is more than 0.2.

3.2 Testing for logistic distribution

Consider logistic d.f. $F(x) = 1/(1 + e^{-x})$, and its density $f(x) = e^x/(1 + e^x)^2$. Note that $\phi(x) = -\dot{f}(x)/f(x) = (e^x - 1)/(e^x + 1)$. With $t = F(x)$, we have

$$\dot{p}(t) = \frac{1 - e^x}{1 + e^x}, \quad \dot{q}(t) = -1 + \frac{x(-1 + e^x)}{1 + e^x}.$$

It is easy to see that

$$\begin{aligned}
\int_{F(x)}^1 \dot{p}(u)^2 du &= \frac{3e^{2x} + 1}{3(e^x + 1)^3}, \\
\int_{F(x)}^1 \dot{p}(u)\dot{q}(u) du &= \frac{1}{3} \ln(1 + e^x) - \frac{e^x \{x(3 + e^{2x}) + (1 + e^x)\}}{3(1 + e^x)^3}, \\
\int_{F(x)}^1 \dot{q}(u)^2 du &= \frac{1}{(1 + e^x)} + \left(-\frac{2}{(1 + e^x)} - \frac{2xe^x}{(1 + e^x)^2} \right) + Re(x),
\end{aligned}$$

where $Re(x) = \int_x^\infty 2s^2 e^s (1 - e^s)^2 / (1 + e^s)^4 ds$. Note that $\int \dot{q}(u)^2 du$ does not have a closed form solution since $Re(x)$ does not have one. However, $Re(x)$ is bounded and decays to 0 fast as x goes to ∞ ; it converges to the finite value (2.43) as x goes to $-\infty$. Therefore, we get a numerical approximation to $Re(x)$ and use it when we calculate the inverse of Γ_F . Let $v_1(x) := \int_{F(x)}^1 \dot{p}(u)\dot{q}(u) du$ and $v_2(x) := \int_{F(x)}^1 \dot{q}^2(u) du$. Also define

$$\begin{aligned}
d(x) &:= \left[3(1 + 3e^{2x})(1 + e^x)^3 v_2(x) - 9(1 + e^x)^6 v_1^2(x) \right]^{-1}, \\
k_1(x) &:= 3(1 + e^x)^3 v_2(x) - 3x(1 + e^x)^3 v_1(x), \\
k_2(x) &:= -3(1 + e^x)^3 v_1(x) + x(1 + 3e^{2x})^3.
\end{aligned}$$

Then, using partitioned four 2×2 blocks of $\Gamma_{F(x)}$ as done in the previous section, we finally obtain

$$\begin{aligned}
& l(\widehat{Z}_i)' \Gamma_{F(x)}^{-1} l(x) f(x) \\
&= 3e^x (1 + e^x) B_{11}(x) \left\{ 1 - 3d(x)e^x \{ (1 - e^x)k_1(x) + k_2(x)(1 + e^x + x(1 - e^x)) \} \right\} \\
&+ \frac{3(1 - e^{\widehat{Z}_i})}{(1 + e^{\widehat{Z}_i})} \left\{ 9e^{2x}(1 + e^x)^2 d(x) B_{11}(x) k_1(x) \right. \\
&+ 3e^x d(x) \{ 3(1 + e^x)^3 (1 - e^x) v_2(x) - 3(1 + e^x)^3 v_1(x) \{ (1 + e^x) + x(1 - e^x) \} \} \\
&+ 27e^{3x}(1 + e^x)^2 B_{11}(x) d^2(x) \{ (1 - e^x)k_1^2(x) + \{ (1 + e^x) + x(1 - e^x) \} k_1(x)k_2(x) \} \} \\
&+ \left(1 + \frac{\widehat{Z}_i(1 - e^{\widehat{Z}_i})}{(1 + e^{\widehat{Z}_i})} \right) \left\{ 9e^{2x}(1 + e^x)^2 d(x) B_{11}(x) k_2(x) \right. \\
&+ 3e^x d(x) \{ -3(1 + e^x)^3 (1 - e^x) v_1(x) + (1 + 3e^{2x}) \{ (1 + e^x) + x(1 - e^x) \} \} \\
&+ 27e^{3x}(1 + e^x)^2 B_{11}(x) d^2(x) \{ (1 - e^x)k_1(x)k_2(x) + \{ (1 + e^x) + x(1 - e^x) \} k_1^2(x) \} \} \Big\},
\end{aligned}$$

and hence, use above quantity for \widehat{U}_n in (2.4).

F	n	KMT	EL1	EL2	KMT	EL1	EL2
logistic	50	0.032	0.076	0.062 (*)	0.020	0.018 (*)	0.004
	100	0.030	0.062	0.060 (*)	0.014 (*)	0.016	0.018
	200	0.026	0.060	0.052 (*)	0.010 (*)	0.020	0.020
	500	0.054	0.052 (*)	0.052 (*)	0.011 (*)	0.006	0.008
normal	50	0.004	0.076	0.080 (*)	0.000	0.022 (*)	0.020
	100	0.020	0.080 (*)	0.066	0.000	0.018	0.020 (*)
	200	0.136 (*)	0.098	0.104	0.040 (*)	0.034	0.028
	500	0.614 (*)	0.176	0.204	0.342 (*)	0.066	0.084
STT	50	0.060	0.083 (*)	0.063	0.040 (*)	0.025	0.019
	100	0.097 (*)	0.070	0.072	0.070 (*)	0.013	0.015
	200	0.132 (*)	0.065	0.063	0.101 (*)	0.019	0.017
	500	0.219 (*)	0.057	0.049	0.173 (*)	0.013	0.013
MTN	50	0.164 (*)	0.089	0.101	0.129 (*)	0.031	0.018
	100	0.185 (*)	0.085	0.097	0.144 (*)	0.029	0.011
	200	0.297 (*)	0.076	0.087	0.234 (*)	0.027	0.028
	500	0.583 (*)	0.115	0.081	0.518 (*)	0.011	0.009
Cauchy	50	0.649	0.965	0.974 (*)	0.517	0.904	0.930 (*)
	100	0.855	1.000 (*)	1.000 (*)	0.782	1.000 (*)	1.000 (*)
	200	1.000	1.000	1.000	1.000	1.000	1.000
	500	1.000	1.000	1.000	1.000	1.000	1.000
Laplace	50	0.137	0.146	0.208 (*)	0.075 (*)	0.025	0.054
	100	0.199	0.350 (*)	0.317	0.159 (*)	0.126	0.093
	200	0.332	0.547 (*)	0.518	0.194	0.320 (*)	0.296
	500	0.613	0.915 (*)	0.911	0.399	0.734	0.754 (*)

Table 3: Empirical level and power obtained from testing $H_0 : F = F_L$

Table 3 reports empirical levels and powers of KMT and EL's for testing $H_0 : F = F_L$; the value corresponding to logistic F stands for the empirical level, and others represent the powers. EL shows better empirical level than KMT for all n 's when $\alpha = 0.05$. With $\alpha = 0.01$, KMT, however, has better one than EL except $n = 50$. Therefore, KMT and EL tie in terms of the empirical level.

For MTN, normal, and STT, the fact that KMT outperforms EL in terms of the power is evident. In case of MTN, KMT overwhelms EL for all n 's and α 's. The difference of powers between KMT and EL increases as n increases and reaches more than 0.5 for both α 's. In the normal case, the maximum power EL attains is 0.204-when $n = 500$ and $\alpha = 0.05$ -while

that of KMT is 0.614. When $n = 500$ and $\alpha = 0.01$, both EL's display powers less than even 0.1 while KMT shows 0.342. When F is STT, EL shows the extremely poor power with $\alpha = 0.01$; EL obtains the power of only 0.013 even though n reaches 500. Note that the counterpart of KMT is 0.173.

When F is either Cauchy or Laplace, EL shows slightly or strictly better power than KMT. For the Cauchy, EL attains the power greater than 0.9 for all α 's even when $n = 50$ while KMT shows the power less than 0.8. Both KMT and EL, however, attain the powers of 1 when n reaches 200; the difference of performances between KMT and EL disappears as n increases. When F is Laplace, EL shows better power than KMT as it does in the previous section. When $n = 500$ and $\alpha = 0.05$, EL attains the power of 0.9; KMT never attains the power greater than 0.8.

4 Conclusion

When we test for normal and logistic distributions, KMT shows better empirical level than EL if α is 0.01 while the opposite is true if α is 0.05. Therefore, it is hard to tell which method is superior in terms of the empirical level. When we test for normality, KMT shows better power than EL if true distribution is logistic, STT, MTN, or Cauchy; EL is superior only if true distribution is Laplace. Similar facts hold when we test for logistic distribution. As stated in the end of section 2, EL's in section 3.1 and 3.2 show very poor powers when true distributions are logistic and STT, respectively.

5 Supplementary material

R-package “GofKmt_1.0.tar.gz”: This package contains a function, “KhmaladzeTrans.”

KhmaladzeTrans provides Khmaladze transformed test statistic. This package is now available at: “cran.r-project.org.”

Acknowledgments.

The author wish to thank his advisor Dr. Hira L. Koul for proposing the topic for this article. The author also expresses gratitude for his generous guidance.

References

- [1] Durbin, J. (1973). *Distribution Theory for Tests Based on the Sample Distribution Function*. SIAM, Philadelphia.
- [2] Fan, J. and Koul, H. (2006) *Frontiers in Statistics*. Imperial College Press, London.
- [3] Khmaladze, E. V. (1979). The use of ω^2 —tests for testing parametric hypotheses. *Theory Probab. Appl.*, **24** 283-301.
- [4] Khmaladze, E. V. (1981). A martingale approach in the theory of goodness-of-fit tests. *Theory Probab. Appl.*, **26** 240-257.
- [5] Khmaladze, E. V. and Koul, H. L. (2004). Martingale transforms of goodness-of-fit tests in regression models. *Ann. Statist.*, **32** 995-1034.
- [6] Peng, H. and Schick, A. (2013). Empirical likelihood approach to goodness of fit testing. *Bernoulli*, **19** 954-981.
- [7] Qin, J. and Lawless, J. (1994). Empirical likelihood and general estimating equations. *Ann. Statist.*, **22** 300-325.
- [8] Stephens, M. A. (1974). EDF statistics for goodness of fit and some comparisons. *J. Amer. Statist. Assoc.*, **69** 730-737.
- [9] Wilks, S. S. (1938) The large-sample distribution of the likelihood ratio for testing composite hypotheses. *Ann. Math. statist.*, **9** 60-2.